Multilevel modelling and public health policy

Leyland, Alastair H.; Groenewegen, Peter P.

Published in:
Scand J Public Health

DOI:
10.1080/14034940210165028

ORIGINAL ARTICLE

Taylor & Francis healthsciences

Multilevel modelling and public health policy

Alastair H. Leyland[1] and Peter P. Groenewegen[2]

[1]MRC Social and Public Health Sciences Unit, Glasgow, Scotland, [2]Nivel Netherlands Institute for Health Services Research, Utrecht, The Netherlands

Scand J Public Health 2003; 31: 267–274

Background: Multilevel modelling is a statistical technique that extends ordinary regression analysis to the situation where the data are hierarchical. Such data form an increasingly common evidence base for public health policy, and as such it is important that policy makers should be aware of this methodology. *Method:* This paper therefore lays out the a basic description of multilevel modelling, discusses the problems of alternative approaches, and details the relevance for public health policy before describing which levels are relevant and illustrating the different kinds of hypotheses that can be tested using multilevel modelling. A series of examples is used throughout the paper. These relate to regional variations in the incidence of heart disease, the allocation of health resources, the relationship between neighbourhood disorder and mental health, the demand-control model in occupational health, and a school intervention to prevent cardiovascular disease.

Key words: multilevel modelling, hypothesis testing, public health policy.

Alastair H. Leyland, MRC Social and Public Health Sciences Unit, University of Glasgow, 4 Lilybank Gardens, Glasgow G12 8RZ, Scotland. Tel: +44 141 357 7504, fax: +44 141 337 2389. E-mail: a.leyland@msoc.mrc.gla.ac.uk

INTRODUCTION

Public health policy is increasingly based on evidence (1). An effective link between policy and research requires an effort from both parties. Research relevant to public health policy increasingly uses a statistical technique called *multilevel analysis* (MLA). This technique is particularly relevant because it enables the analysis of one of the ''articles of faith'' of public health: that public health is about the health of people in their context (2).

Pioneering development of MLA methodology has been in education where researchers have been interested in studies examining how pupil outcomes (such as examination scores) are related to the characteristics both of the pupils themselves and those of the schools (3). The use of MLA has since been widespread in the overlapping fields of health services research, epidemiology, and public health (4–7), assisted by the development of specialist multilevel software and the addition of multilevel capabilities to common statistical packages (8). The educational example may be transferred to a public health context in several ways. For example, interest focuses on the roles played by the hospital and patient when studying patient outcomes. The individual and the workplace may both influence absence from work due to sickness. Regional differences in incidence of heart disease may reflect differences in the composition of populations and in the success of local health promotion programmes.

In this article we shall clarify what MLA is about and why it is becoming popular in public health research. This popularity makes it important for public health professionals to have a basic understanding of this technique so as to be able to judge research that uses it.

We shall start by explaining what MLA is and how it relates to single-level regression analysis. We shall explain how, using MLA, we try to avoid both the ecological and individualistic fallacies. We will discuss the relevance for public health of being able to analyse variables from different levels simultaneously. We will expound on the idea of a level by discussing what a level is and which levels are particularly relevant in public health.

WHAT IS MULTILEVEL MODELLING?

We consider regional variation in coronary heart disease (CHD) as an example to illustrate what MLA is. CHD incidence varies across individuals – some, but not all, people have heart disease. The incidence rate also varies across regions; this suggests that some of the variation in CHD incidence is at the level of the individual and some at the level of the region. These

 ISSN 0301-7311

DOI: 10.1080/14034940210165028

are the levels in our model of CHD incidence. Individual characteristics may explain some individual variation; incidence may be lower among women, may increase with age and may be highest among blue-collar workers. Adjusting for individual characteristics may reduce the variation between *regions* as well as that between individuals. If a relatively high proportion of blue-collar workers live in one region then it is likely to have higher incidence than those regions with high proportions of white-collar workers. Such standardization forms the basis of any secure comparison. However, how do we proceed to explain any remaining differences between the standardized incidence rate of regions? Just as the variation between individuals may be (partly) explained by their characteristics, so we may use characteristics of the regions to explain some of the differences between regions. For example, we may consider whether the regional public health authority had introduced a specific health promotion programme or the proportion of the region's GDP spent on healthcare. Such factors may explain some variation between regions but will obviously not affect the variation between individuals within regions. There is a large body of literature describing how context may affect individual health (9–14).

MLA lets us separate or partition the total variation in the incidence of CHD into that part due to differences between individuals and that due to differences between regions. It also means we can explore those characteristics of individuals or regions that may explain these differences.

MLA is an extension of ordinary least squares (OLS) regression analysis under which we may estimate, for example, the mean relationship between CHD incidence and age in our population, assuming that there is no regional effect above the characteristics of the individual. Such a relationship is illustrated in Fig. 1a. The algebraic notation of an OLS regression equation is:

$$y_i = \beta_0 + \beta_1 x_i + e_i$$

where β_0 is the intercept of the regression line (the value of the dependent variable y taken when $x = 0$), β_1 is the slope associated with the independent variable x and e_i is the residual for the i^{th} individual.

An advantage of MLA is indicated in an alternative name, *random coefficient* modelling. Random coefficients come in at two points. First, the average outcome (mean incidence of CHD) for each region may differ; this mean is modelled as a random sample from a hypothetical distribution of all possible regions. The relationship between CHD and age is assumed to be the same in all regions; what we are really fitting is a set of parallel lines indicating that the mean incidence of CHD differs between regions for patients with a certain set of characteristics (e.g. the same age). Fig. 1b illustrates a random intercept model. Algebraically:

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_j + e_{ij}$$

where u_j is the residual of the higher level unit (region) and e_{ij} the residual associated with an individual within this higher level unit. The higher level residual u_j is an effect of the j^{th} region shared by all individuals within that region.

Second, the relationship between individual characteristics and susceptibility to CHD may differ between regions. In some regions the increase in CHD incidence with age may be more pronounced. To take account of such differential relationships the regression slopes are allowed to differ between regions and again these slopes are modelled as a random sample. Such a model is shown in Fig. 1c.

Fig. 1. Relationship between incidence and age (*a*) where there is no regional effect; (*b*) where the regional effect is independent of age and (*c*) where the regional effect varies with age.

Algebraically:

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_{0j} + u_{1j} x_{ij} + e_{ij}$$

where u_{1j} is the slope residual in region j just as u_{0j} is the intercept residual.

WHY USE MULTILEVEL ANALYSIS?

Before MLA we had two options when analysing data from different levels with OLS regression. We could either aggregate the data to the higher level or we could distribute the characteristics of higher levels to all individuals. Both options bring their problems. Aggregating and distributing the data pose problems of interpretation of the results, known as the *ecological* and the *atomistic* fallacy respectively. Distributing the data gives an additional problem with the estimates of the higher level characteristics. We will discuss each of these problems.

To illustrate the idea of aggregation we could use the regional incidence rate of CHD as the dependent variable and variables including average age and income, proportion of women etc. as independent variables. This loses a lot of information and we risk misinterpreting the results – the ecological fallacy. This is the methodological identification of a relationship at an area level between an outcome and a population characteristic, and attribution of this relation to individuals when this relationship actually does not exist at the individual level.

The alternative is to distribute regional characteristics to all individuals. For example, the economic welfare of regions would be assigned to all individuals and would be identical for each individual within a region. Here we risk the atomistic fallacy – the methodological identification of a relationship between an outcome and an individual characteristic, and attribution to the context level, when this relationship does not exist if the context of individuals is taken into account or when the relationship varies between contexts (15).

An illustration of the ecological fallacy was given in the context of resource allocation (16). Here we expand upon that example to illustrate also the atomistic fallacy.

When considering the allocation of resources for healthcare to municipalities a common approach is to use the relationship between previous utilization and measures of need to predict future utilization (17). Assume that the data available for each individual reflect a socioeconomic indicator that is indicative of the need for healthcare – perhaps a composite score encompassing factors such as income, employment, and living circumstances – and the actual cost of the delivery of care to that person. (In practice such data may only be available for small areas rather than individuals, but the same arguments hold if those areas form part of the municipalities with responsibility for healthcare for their populations.) Our problem then simplifies to identifying the relationship between cost and need, so that resources may be allocated to municipalities on the basis of current need to cover future costs.

Fig. 2a illustrates the relationship between cost and need found among individuals. As need increases, so does the average cost. We have ignored the fact that these individuals live in different municipalities and have conducted our analysis without considering their context. Fig. 2b shows the relationship between cost and need across three municipalities. The relationship differs little from that found at the individual level (Fig. 2a). This ignores the data on individuals and assumes that the average relationships between municipalities hold between individuals. Fig. 2c shows the full picture. The relationship between cost and need between individuals is fairly consistent

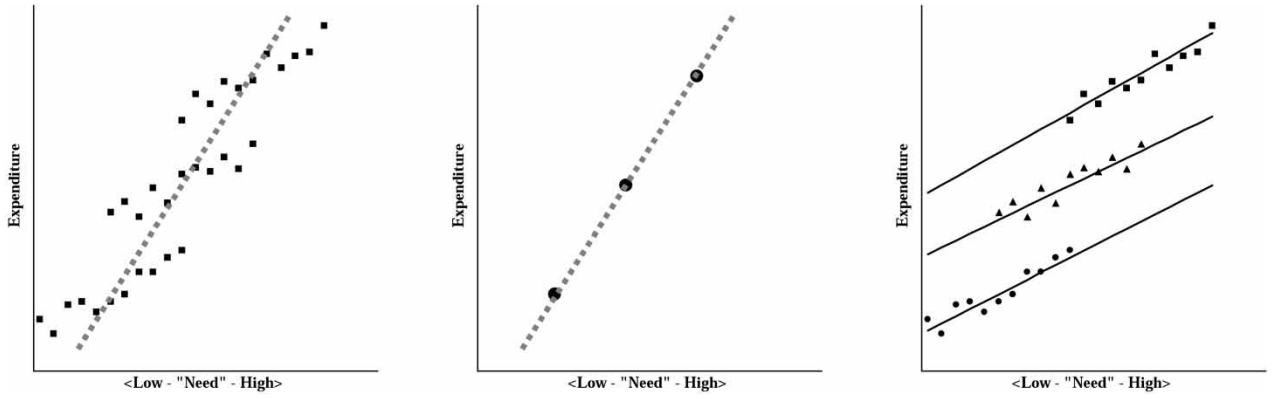

Fig. 2. The relationship between expenditure and need based on (*a*) individual data; (*b*) data aggregated to the area level and (*c*) the multilevel data structure being taken into account.

across the three municipalities, with an increase in need being associated with an increase in cost, but the slope is not as steep. An increase in need is associated with a smaller increase in cost than in Fig. 2a or Fig. 2b. This is because the average level of spending for a fixed level of need varies between municipalities, and the ecological and individual analyses could not take this into account. Only MLA uncovered the true relationship between need and cost.

Apart from the atomistic fallacy, distributing data to lower levels is associated with a statistical problem in determining the significance of relationships. The precision of any estimate is determined, in part, by the sample size – the larger our sample, the more precise our estimate. However, if the outcome varies across high-level units then individuals cannot be considered to be independent. This has the effect of reducing the effective sample size – in an extreme case, if all individuals in the same region respond in the same way then our sample size would just be the number of regions in the study – and, consequently, of decreasing the precision of our estimate. In practical terms the consequences of ignoring the multilevel data structure is that too many relationships will be found to be "significant". This problem is known as *misestimated precision*. The extent to which individuals in the same region respond in the same way is quantified by the *intraclass correlation* (ICC), defined as the higher level variation divided by the total variation in the outcome variable. The ICC tells us how strong the resemblance is between lower level units nested within the same higher level unit. A priori estimates of the ICC can be used to choose the optimal sampling design in terms of the numbers of units at different levels.

RELEVANCE FOR PUBLIC HEALTH RESEARCH AND POLICY

Public health policy relates to different areas. We have illustrated the role that MLA might play in health administration and planning through the example regarding resource allocation. Other areas include community health and health protection, occupational health, and community-level interventions. We will briefly discuss each of these broad areas, indicating the contribution of MLA in analysing research problems with examples.

Community health and health protection deals with people in their social and environmental context. People's health is not only determined by their personal characteristics and their biological and genetic endowment. It is also influenced by their exposure to environmental influences of a physical kind, to stressing circumstances, and to the influence of family or neighbours (18). Some influences are direct, others work by influencing people's behaviour or coping strategies. When studying people in context, different levels are always involved: individuals as the lower level and a specific context as the higher level. These contexts depend on the research problem and may range from the family to the neighbourhood or to areas that are defined by equal extent of air pollution. (The next section considers which contexts are relevant in a given research problem.)

MLA makes it possible to test three kinds of hypotheses: first, the hypothesis that only individual characteristics are responsible for health differences between communities. If individual characteristics related to health cluster in some communities, one might mistake this for differences produced by community characteristics or circumstances. For example, some communities may have poorer health outcomes but at the same time have older populations. MLA makes it possible to distinguish these so-called *compositional* effects from real contextual effects. One could pose the question as to why people with certain characteristics cluster together. The identification of compositional effects does not solve the problem of individual choice versus material conditions.

Second, if there are contextual effects, MLA enables hypothesis testing about the relationship between contextual characteristics and health, taking individual influences on health into account. This provides better estimates of the relation between context and health. An example would be analysing the effect of community wealth on population health, taking individual income into account. Finally, MLA makes it possible to study specific combinations of individual and contextual characteristics, so-called *cross-level interactions*. As an example one could hypothesize that people on low incomes in high-income neighbourhoods have poor health status (relative to people on low incomes in *low*-income neighbourhoods). The interpretation would be that the absolute level of wealth is not that important for people's health but relative deprivation is. This would be modelled using an interaction between average and individual income.

An example of a study relevant for community health concerns the relationship between neighbourhood disorder and the mental health of the residents (19). The reasoning is that the neighbourhood social context leads to neighbourhood disorder (level 2), which is hypothesized to influence the mental health of the residents (level 1). This relationship is studied, taking into account individual characteristics known to correlate with mental health (Fig. 3).

The second example concerns occupational health.

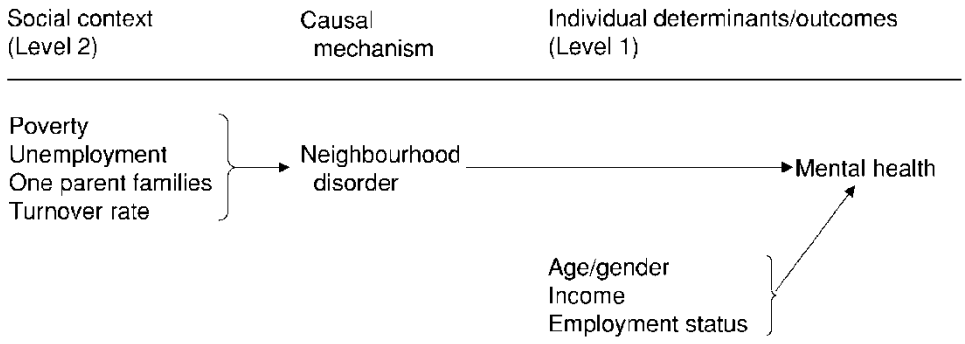

Fig. 3. Neighbourhood disadvantage and depression.

An often used model to explain variations in people's health in relation to their work situation is the demand-control model (20–21). The health of the working population is influenced by both their individual characteristics, such as susceptibility to illness (level 1), and workplace-related characteristics (level 2). The demand-control model hypothesizes that psychological job strain, leading to poor health, is the joint effect of the demands of a work situation and the control people have over their own work. These characteristics of the work situation are based on the organization of work (Fig. 4).

Our third example concerns group or community health interventions. The reasons for delivering an intervention to a community rather than to individuals include the reduced cost (for example, for an educational programme), the impracticality or impossibility of conducting an individual-level intervention (for example, for a water fluoridation programme), and because of the inability to avoid cross-contamination at an individual level (for example, an educational programme where intervention and control groups may meet). The unit of randomization to intervention or control may be, for example, the school, workplace, area of residence, or general practice, depending on the delivery of the intervention and the nature of the study (22). An example of such a trial is the child and adolescent trial for cardiovascular health (CATCH) (23), a programme aimed at school-children in four states of the USA with the aim of primary prevention of cardiovascular disease. Outcomes for individual children (serum cholesterol levels and psychosocial factors) may be influenced not only by individual characteristics such as age, height, and body mass index. The trial sought to ascertain whether they could also be influenced by the fat content of lunches and the amount of exercise taken, both of which could be modified by intervention at school level (Fig. 5).

A common element in these examples is that the outcomes are at the individual level and are the joint product of individual-level characteristics and characteristics at the level of a social context. Public health research is full of such relationships and MLA allows us to analyse them in an appropriate way.

WHAT IS A LEVEL AND WHICH LEVELS ARE RELEVANT?

First of all we need to discuss what constitutes a level. Earlier we introduced the notion of random coefficients; CHD incidence was used as an example. Average CHD incidence varied between regions and was modelled as a random sample of incidence rates from some hypothetical distribution. This way of modelling the incidence of CHD supposes that the

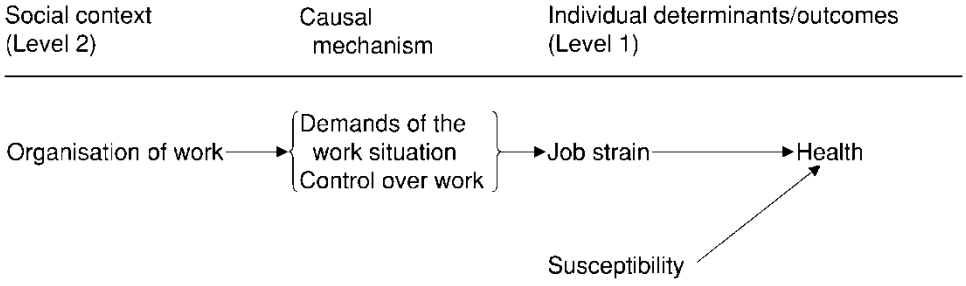

Fig. 4. The demand-control model of occupational health.

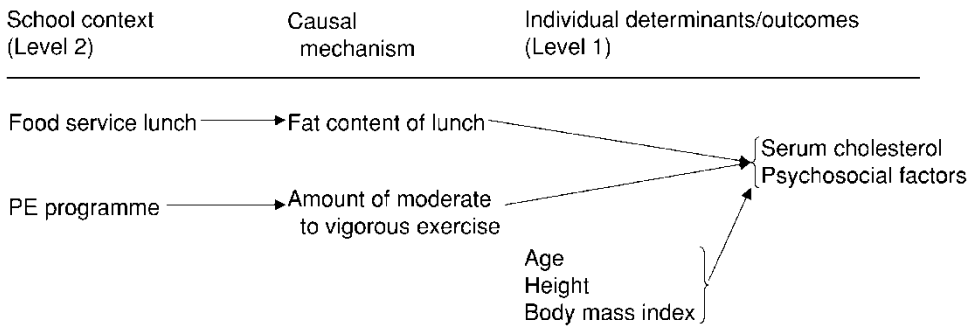

Fig. 5. Community (school) intervention.

higher level consists of units that can be meaningfully sampled. Here that would be a sample of regions from a population of regions. In practice we often work with *all* regions rather than a sample; in such a situation these can be considered a sample for the generalizability of results. The data for each region form a sample of the data that could possibly have been collected at different times and allow us to make inferences about those regions and regions in general.

Generally, levels comprise units that can be observed, sampled, and analysed. These units have characteristics that can either be directly observed and measured, such as traffic density in a neighbourhood, or aggregated from individual characteristics, such as average income.

The distinction between a level and its characteristics is important. Any characteristic, such as the degree of urbanization of regions, may have a number of values – e.g. in six classes from highly urban to sparsely populated countryside. Categories of urbanization are not something that we usually sample. We do sample, however, municipalities or regions and categorize them according to urbanization. Urbanization is a variable and municipalities are units that, among other things, can be characterized by their degree of urbanization.

In survey research urbanization can both be used at the individual or municipality level depending on the sampling design. In health interviews people are asked questions about health-related behaviour and subjective health. Characteristics of the place where people live may also be requested or recorded. The dataset than comprises details about the individuals interviewed and a variable concerning the place where they live. It is possible to study the relationship between degree of urbanization and (e.g.) mental health. All units are still at the individual level; there is no sampling of municipalities. Alternatively, the sample design of the same health interview survey could be such that, first, a number of municipalities are sampled and, within each of the sampled municipalities, a sample of interviewees is drawn. The dataset now contains individual data and the identity of the municipality. Characteristics of the municipality can be added from other sources or constructed by aggregating individual variables. The result is a database with sampled units at two levels.

In survey practice, a simple random sample is rarely drawn for pragmatic reasons – consider the costs of interviewing people over the whole country. Usually a staged sample is used but often it is treated as a simple random sample or simplistic adjustments are made for the sample design. With the diffusion of MLA in health-related research, there are now tools to treat a two-stage sample in an appropriate way and it has become more common to theorize about the way context affects people's health, health-related behaviour, and health service utilization.

Higher level units are important because they define the action space of individuals. Many problems in public health policy are related to people's behaviour (see the earlier examples). People behave within the social and institutional context of their community or workplace. This context influences the resources and the range of options (opportunities and constraints) that actors have (24).

The question "Which levels are relevant?" is answered by analysing the research problem and asking: "What kind of opportunities and constraints determine people's behaviour and in which units are these opportunities and constraints patterned?" This abstract notion is illustrated with an example.

When considering neighbourhood differences in health there are at least three constraints that influence people's health or health behaviour:

1. People live in social units and these offer opportunities and constraints that influence people's health and health behaviour. Neighbourhoods differ in how close the relations between people are and the availability of support networks. Social integration and social support are known to influence people's health. A relevant level would comprise small-scale, socially homogeneous units.

2. People also live in administrative and planning areas, which are used to plan and organize healthcare facilities, including community health centres and hospitals, and to organize public health activities, such as anti-smoking campaigns. Here the opportunities and constraints are more institutional.
3. People's health is also influenced by exposure to the physical environment. Areas differ in exposure to factors including noise and air pollution. When analysing these physical influences units at a level higher than socially homogeneous units or administrative units are relevant (25).

Different constraints related to several higher levels could influence health at the same time, either separately or jointly. Different levels work in conjunction if for example municipalities have certain policies and their effectiveness depends on the characteristics of neighbourhoods within these municipalities.

In conclusion, examples of higher level units relevant in public health research are administrative areas, such as municipalities; social units, such as groups of neighbours or peers; service areas of healthcare institutions, such as hospitals; places of work, such as different departments of a large enterprise; and exposure regions. Ideally, the choice of a higher level should not depend on what routinely collected administrative data are available but on a substantial analysis of the research problem. However, for practical reasons one often has to accept data on administrative units, while actually needing data based on areas with different levels of exposure. In these sub-optimal cases it is important, when interpreting results, to be aware that the units of interest may not be exactly those that have been sampled.

CONCLUSIONS

MLA is a powerful tool for analysing behaviour and its consequences in context. That makes it important for public health research, health services research and epidemiology. MLA allows us to split variation into that part related to individuals and that related to higher level units. By formulating and testing hypotheses we seek to explain variation at these levels.

The higher level units must be relevant for our understanding of the outcomes studied. Classical public health looked for social and economic influences on the population's health (poverty, public hygiene, housing, regulation of labour etc.). These social and economic influences are patterned in different higher level units. Through the development of modern epidemiology as a research tool in public health, the emphasis on context has probably diminished. Modern epidemiology has contributed enormously to our understanding of the aetiology of ill health but, compared with classical public health, the social, economic, and population perspectives have been more in the background. MLA provides the tools for a new integration of these approaches and perspectives.

MLA is increasingly used in public health research. Whenever data contain different levels, MLA is indicated. The question is not so much whether but how it should be used. It might be argued that in cases where the effect of context on health behaviour or outcomes is negligible, MLA is irrelevant. But how are we supposed to know? The relative importance of context and individual is an empirical question and unknown beforehand. That alone is reason to use MLA, if only to show it to be unnecessary in particular cases. In this paper we have explained the relevance of MLA for public health research. Both researchers and policy makers in public health need to be aware of these methods. An increasing number of research papers use them and it is important that the readership of these papers has sufficient understanding of the background of multilevel research.

ACKNOWLEDGEMENT

This article is based on a lecture given by the authors to public health administrators in Scania council during a multilevel research course organized by the Swedish Society of Social Medicine in October 2000. The Social and Public Health Sciences Unit is supported by the Medical Research Council and the Chief Scientist Office of the Scottish Executive Health Department.

REFERENCES

1. Rosen R. Applying research to health care policy and practice: medical and managerial views on effectiveness and the role of research. J Health Services Res Pol 2000; 5: 103–8.
2. Rosen G. A history of public health (1958). Baltimore: Johns Hopkins University Press, 1993.
3. Snijders TAB, Bosker RJ. Multilevel analysis: an introduction to basic and advanced multilevel modeling. London: Sage, 1999.
4. Rice N, Leyland AH. Multilevel models: applications to health data. J Health Services Res Pol 1996; 1: 154–64.
5. Diez-Roux AV. Multilevel analysis in public health research. Annu Rev Public Health 2000; 21: 171–92.
6. Keskimäki I, Karvonen S, Sund R, Leyland AH. Monitasomallien käyttö terveystutkimuksessa (Multilevel modelling in health research). Sosiaalilääketieteellinen Aikakauslehti (J Soc Med) 2001; 38: 327–35.
7. Leyland AH, Goldstein H, editors. Multilevel modelling of health statistics. Chichester: Wiley, 2001.
8. de Leeuw J, Kreft IGG. Software for multilevel analysis.

In Leyland AH, Goldstein H, editors. Multilevel modelling of health statistics. Chichester: Wiley, 2001: 187–204.
9. Duncan C, Jones K, Moon G. Do places matter: a multilevel analysis of regional variations in health related behaviour in Britain. Soc Sci Med 1993; 37: 725–33.
10. Sloggett A, Joshi H. Higher mortality in deprived areas: community or personal disadvantage? Br Med J 1994; 309: 1470–4.
11. Diez-Roux AV, Nieto FJ, Muntaner C, Tyroler HA, Comstock GW, Shahar E, Cooper LS, Watson RL, Szklo M. Neighbourhood environments and coronary heart disease: a multilevel analysis. Am J Epidemiol 1997; 146: 48–63.
12. Blakely T, Woodward A. Ecological effects in multilevel studies. J Epidemiol Community Health 2000; 54: 367–74.
13. Lynch J, Davey Smith G, Kaplan G, House J. Income inequality and mortality: importance to health of individual income, psychosocial environment, or material conditions. Br Med J 2000; 320: 1200–4.
14. Macintyre S, Ellaway A, Cummins S. Place effects on health: how can we conceptualise, operationalise and measure them? Soc Sci Med 2002; 55: 125–139.
15. Duncan C, Jones K, Moon G. Context, composition and heterogeneity: using multilevel models in health research. Soc Sci Med 1998; 46: 97–117.
16. Smith PC, Rice N, Carr-Hill RA. Capitation funding in the public sector. J Roy Stat Soc Series A 2001; 164: 217–57.
17. Carr-Hill RA, Sheldon TA, Smith PC, Martin S, Peackock S, Hardman G. Allocating resources to health authorities: development of methods for small area analysis of inpatient services. Br Med J 1994; 309: 1046–9.
18. Macintyre S, MacIver S, Sooman A. Area, class and health: should we be focussing on places or people? J Soc Pol 1993; 22: 213–34.
19. Ross CE. Neighborhood disadvantage and adult depression. J Health Soc Behav 2000; 41: 177–87.
20. Söderfeldt B, Söderfeldt M, Jones K, O'Campo P, Muntaner C, Ohlson CshyG, Warg LshyE. Does organization matter? A multilevel analysis of the demand-control model applied to human services. Soc Sci Med 1997; 44: 527–34.
21. Van Yperen NW, Snijders TAB. A multilevel analysis of the demands-control model: is stress at work determined by factors at the group level or the individual level? J Occup Health Psychol 2000; 5: 182–90.
22. Koepsell TD, Wagner EH, Cheadle AC, Patrick DL, Martin DC, Diehr PH, Perrin EB, Kristal AR, Allan-Andrilla CH, Dey LJ. Selected methodological issues in evaluating community-based health promotion and disease prevention programs. Annu Rev Public Health 1992; 13: 31–57.
23. Luepker RV, Perry CL, McKinlay SM, Nader PR, Parcel GS, Stone EJ, et al. Outcomes of a field trial to improve children's dietary patterns and physical activity. JAMA 1996; 275: 768–776.
24. Groenewegen PP. Dealing with micro–macro relations: a heuristic approach with examples from health services research. In: Westert GP, Verhoeff RN, editors. Places and people: multilevel modelling in geographical research. Utrecht: Nederlandse Geografische Studies, 1997: 9–18.
25. Dunn CE, Kingham SP, Rowlingson B, Bhopal RS, Cockings S, Foy CJW, Acquilla SD, Halpin J, Diggle P, Walker D. Analysing spatially referenced public health data: a comparison of three methodological approaches. Health and Place 2001; 7: 1–12.

Accepted 02 10 10